DOCUMENT RESUME

ED 476 967 IR 021 695

AUTHOR	Bain, John D.; Mavor, Ken
TITLE	Collaborative Teachback with a Statistical Cognitive Tool: A Formative Evaluation.
PUB DATE	2002-06-00
NOTE	8p.; In: ED-MEDIA 2002 World Conference on Educational Multimedia, Hypermedia & Telecommunications. Proceedings (14th, Denver, Colorado, June 24-29, 2002); see IR 021 687.
AVAILABLE FROM	Association for the Advancement of Computing in Education (AACE), P.O. Box 3728, Norfolk, VA 23514. Tel: 757-623-7588; e-mail: info@aace.org; Web site: http://www.aace.org/DL/.
PUB TYPE	Reports - Research (143) -- Speeches/Meeting Papers (150)
EDRS PRICE	EDRS Price MF01/PC01 Plus Postage.
DESCRIPTORS	Cognitive Structures; Computer Oriented Programs; *Cooperative Learning; *Educational Technology; Formative Evaluation; *Least Squares Statistics; Small Group Instruction; Statistical Data; *Statistics; *Student Developed Materials

ABSTRACT

A learning environment is described in which students collaborate in small groups to develop screen movies in which they use a statistical cognitive tool to interpret published research and to demonstrate their understanding of least squares statistical concepts. Evaluation data are reported, which indicate that, although some groups thrive in this environment, others struggle to cope. Enhancements are proposed based on the outcome of the evaluation. Highlights include: the computer program; the course context; the learning process; the intended learning outcomes; the evaluation; and future development. (Contains 20 references and 3 tables.) (Author/AEF)

Reproductions supplied by EDRS are the best that can be made
from the original document.

ED 476 967

Collaborative Teachback with a Statistical Cognitive Tool: A Formative Evaluation

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY

G.H. Marks

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

1

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

This document has been reproduced as received from the person or organization originating it.

Minor changes have been made to improve reproduction quality.

Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

By: John D. Bain & Ken Mavor

BEST COPY AVAILABLE

IR021695

Collaborative Teachback with a Statistical Cognitive Tool: A Formative Evaluation

John D Bain
School of Curriculum, Teaching and Learning, Griffith University, Australia. Email: J.Bain@mailbox.gu.edu.au

Ken Mavor
Department of Psychology, University of Southern Queensland, Australia. Email: mavork@usq.edu.au

Abstract: A learning environment is described in which students collaborate in small groups to develop screen movies in which they use a statistical cognitive tool to interpret published research and to demonstrate their understanding of least squares statistical concepts. Evaluation data are reported which indicate that, although some groups thrive in this environment, other struggle to cope. Enhancements are proposed based on the outcome of the evaluation.

The teaching of applied statistics to social science undergraduates is a much discussed topic (Becker, 1996), no doubt because it presents a disproportionate challenge to students, many of whom do not have extensive mathematical preparation. Although various teaching methods have been adopted, many use prepared data sets in which relationships between variables are to be revealed (and tested for significance) through analyses undertaken by the student (****). In recent years the data sets have been related to research-based scenarios (e.g., Derry et al., 1995; Fischer, 1996; Thompson, 1994), and the analyses are performed with one of a growing number of computer packages which have user-friendly interfaces and powerful graphing capabilities (e.g., *DataDesk, SAS/JMP, SPSS, Statistica, Models'n'Data, Statview*).

The approach upon which the present work is centred shares several of these features (exercises anchored to published research, use of a computer program in the learning process, heavy emphasis on graphical representation), but it differs in several important respects:

- the data are created and altered by the student using a graphical rather than spreadsheet interface
- the emphasis is upon modelling—and playing 'what if' with—the relationships between the variables represented in the program
- the students' task is to demonstrate an understanding of statistical and research concepts by recording a screen movie, with voice-over, in which they 'teachback' what the statistical and research concepts mean and why they are relevant to the research scenario
- the students work in small groups (2–4) to develop and record their teachback movies.

This paper provides details of the teaching/ learning arrangements for this approach to statistics, and reports a formative evaluation of the learning process and outcome.

The Computer Program

The program upon which the learning process is centred is *BivarDescribe*, a Macintosh program designed by the first author. This program enables users to explore the least squares properties of correlation, regression and one-way analysis of variance (anova). Consistent with calls for regression and anova to be taught as variants of the general linear model (Thompson, 1993), anova is treated as a special case of multiple regression, and the focus is upon how the relational structure of the data is modelled by the least squares approach, not on sampling distributions and statistical significance. Thus the program is intended to redress some pervasive misconceptions, namely: that anova and correlation are different procedures (Keppel & Zedeck, 1989; Thompson, 1993); that the concept of 'explained variance' (or 'predictable variance') only applies to correlational data (Hays, 1981; Huberty, 1987); that statistical significance is more important to research interpretation than effect size (Rosenthal & Rosnow, 1985; Thompson, 1993; Wilkinson et al., 1999); and that omnibus anova adequately models relational structure in a one-way design (Rosnow & Rosenthal 1989; Thompson, 1993).

As already noted, this program differs from statistical analysis programs because it facilitates the creation of data (by clicking the 'dot' cursor on the data plot) and the alteration of data (by dragging data points to new locations, or altering variances with a 'variance change' tool). An example of the main plot window in which a data point has been dragged to an outlying position is provided in Figure 1.

Figure 1. The main plot window of *BivarDescribe* depicting a positive linear relationship between two variables, *Severity of Eating Problem* and *Neuroticism*, but with one data point dragged to an outlying position.

Some of the additional features of this program are:

- the strength of the relationship or effect size (the Proportion of Predictable Variance) is represented by a vertical meter that changes dynamically as points are added or altered
- users may access additional information about applicable statistics by opening mini-windows in which the statistics are defined algebraically and illustrated numerically
- there is a dynamic coupling between the graphical and numerical states of the program (points selected or changed in one are selected or changed in the other)
- users may use the Enquiry Tool to see how scores are predicted or to view the deviations with which predicted and residual variances are calculated
- the user can specify a linear prediction and compare it with the least squares regression line
- a correlational design can be converted to an anova design
- contrasts between the conditions of an anova design can be specified graphically and turned on and off to see how they contribute to the predictable variance
- anova as a special cases of regression can be seen in additional window which shows the prediction of scores from the weights defining each contrast.

The Course Context

The course in which the present project is embedded is a one semester second year course in research methods for psychology students at The University of Southern Queensland. The content includes a wide range of research design and methods concepts, for which the supporting text is Shaughnessy, Zechmeister, & Zechmeister, (2000), as well as the statistical concepts involved in bivariate regression, multiple regression and analysis of variance, for which the supporting text is Keppel & Zedeck (1989). In addition, there are extensive course notes (the course is also offered in external mode) and lectures and tutorials. The exercises with which *BivarDescribe* is used comprise 25% of the assessment for the course. Students are prepared for the exercises during several tutorials in one of which they develop and rehearse a simple teachback movie. Students also have access to demonstration movies made with *BivarDescribe*

by the authors (about 8 hours in total duration). These movies are based on a research study not included in the exercises, and their topic structures differ from those required by the exercises (to avoid direct modelling).

The Learning Process

The intended learning process consists of several interdependent activities:

- The 'priming' of relevant statistical and research concepts by viewing the demonstration movies and using them as analogues, not direct models, for the exercise movies.
- The construction of statistical understanding using *BivarDescribe* as a cognitive tool with which "[l]earners themselves function as designers using technologies as tools for analysing the world, accessing information, interpreting and organizing their personal knowledge, and representing what they know to others." (Jonassen & Reeves, 1996, p. 694).
- The mapping back and forth between real research variables and relationships and their representations in the statistical domain (Derry, Levin & Schauble, 1995; Laurillard, 1993)
- Collaboration in small groups to encourage the development and refinement of understanding (Roschelle, 1992)

The Intended Learning Outcomes

The learning outcomes of most relevance here are focused on the ability of students to understand and use least squares statistical methods to interpret data within a realistic research setting, as manifested in the teachback movies created by each group. 'Understanding' in this context has several aspects: (a) the ability to map a research question into a statistical framework and vice versa; (b) the ability to translate flexibly between the graphical, algebraic and arithmetical representations of statistical relationships with a clear sense of why they correspond and what they mean (i.e., the understanding should be 'relational' rather than 'instrumental' —Skemp, 1976); (c) the ability to explain how basic least squares concepts apply to correlation and anova (Thompson, 1993); and (d) the ability to 'perform' understanding rather than to reproduce it (Perkins & Blythe, 1994) as a consequence of the 'teachback' format (Pask, 1976).

Evaluation

The aim of the evaluation reported here was not to determine whether the learning model outlined above is better than others reported in the literature, but rather to determine whether it is functioning much as intended and to suggest improvements (Bain, 1999). Accordingly, data were collected about the learning process (by video-recording the discussions of participating groups and coding for key features of the process) and about the learning outcome (by coding the teachback movies submitted by each group).

Participants

Eighteen groups (54 students) completed the assessment for the course, of which 10 groups agreed to take part in the evaluation (i.e., agreed to have their discussions video-taped). Four of the ten participating groups dropped out of the evaluation early, leaving 6 groups from which the present data were obtained. Most groups comprised 3 students, the range being from 2 to 4.

Teachback Exercise

Two substantial exercises comprised the assessment for the teachback component of the course, for each of which several teachback movies were to be produced. The data reported here are for the first of four movies in the first exercise for which the research scenario was based on the article by McFarland (1989). The scenario asked students to contemplate a study examining the relationship between religious orientation (as measured by the *Quest* self-report scale) and a scale measuring a general tendency to discriminate (*Discrimination*) consisting of attitudes toward several minority groups. The expected relationship was inverse, given the findings reported by McFarland--i.e., the higher the

score on *Quest*, the lower the *Discrimination* score—but this was not stated in the scenario outline. The first movie was specified as follows:

Use BivarDescribe to model the data pattern (relationship) predicted in the scenario, showing the appropriate variable labels and scales. Initially assume a fairly weak relationship, and show what that might look like. Then:

- *refer to the Proportion of Predictable Variance (PPV) indicator to describe and explain the strength of the relationship that you have created*
- *open the PPV window and explain what the predicted scores are and what the proportion of predictable variance means*
- *apply the Variance Change tool to the grand mean (using increasing as well as decreasing modes) and describe and explain (a) the changes in the data and (b) what happens to the Least Squares Coefficient (LSC) and PPV indicators*
- *make sure you tie all your explanations back to the scenario study and discuss the meaning of the data patterns in terms of the scenario variables.*

Procedure

Our intention was to record two learning sessions, one early and one late in each group's preparations for an exercise so that improvements during learning could be documented. However, for logistical reasons this proved impossible to arrange, so the data reported below are for one session conducted part way through the preparation for the exercise. The session was recorded with a split screen format that enabled the group members to be seen as well as the details of activity on the computer screen. The teachback movies comprising an exercise were recorded when the group indicated that it was ready to do so, typically 2-3 weeks after preparations began. The recording was managed by a tutor who operated the recording software (*Snapz Pro 2*), leaving the group free to concentrate on producing its teachback movies. The video recordings and teachback movies were coded by a project assistant using a 5 point scale ranging from 'not at all' through 'poor', 'limited' and 'adequate' to 'excellent', where the descriptors for each point on the scale were contextualised to the characteristic being rated. The coding process involved repeated checks with the senior author about the codes assigned to each case.

Findings

The means reported in the 2nd last column of Table 1 indicate that the groups were better able to effect the intended social aspects of the learning process (discussion and collaboration) than they were able to use the resources to *construct* their understanding of statistical concepts (building understanding with *BivarDescribe* and correcting misunderstandings being rather limited). However, these general trends are qualified by substantial differences between the groups, two (F, G) being conspicuously sound, two adequate (E, J) and two others obviously struggling to cope (D, K). Similar patterns were evident when general and specific aspects of understanding were coded (Table 2). In this case it was possible to use the same scales to code understanding during the learning processes as well as in the teachback movies, and the mean data and ranges are reported in Table 2. Although the mean values on both sets of scales were mostly in the 'limited' to 'adequate' range, this pattern masks the fact that there were marked differences between the groups much as was evident in the process scales: the ranges provide some idea of the performance differences involved, although not the consistent patterns of group performance (space limitations prevent reporting of the group data).

Another aspect of the consistency of group performance is the degree of correspondence between the understanding 'profiles' (pattern of scores on the understanding scales) obtained during the learning process and evident in the learning outcome. An index of profile similarity is provided by the Euclidean Distance measure (Table 3) which varies from zero when the profiles are identical to an empirical maximum when the profiles are most dissimilar. As is evident from Table 3, the understanding profiles of the groups were relatively similar between the preparation sessions and the teachback movies. In other words, the level of understanding reached about mid-way through group discussion was similar to the understanding evident in the final product. This suggests that further assistance is needed to ensure that knowledge construction continues to grow as groups discuss their approaches to the teachback movies.

Future development

The evaluation data reported here (and supported by the other data collected) indicate that some groups were able to construct their understanding in the intended manner, two groups in particular being conspicuously capable, but at least two groups (and maybe also those groups which withdrew early in the semester) were not able to perform to a satisfactory standard. Three enhancements may assist such groups in the next offering of the course:

- there will be more extensive lead-up work in the lab to encourage greater fluency with the software and statistical concepts before the teachback exercises begin in earnest;
- a more useful retrieval and playback interface for the demonstration movies will be available: statistical and research concepts and associated movies will be accessible with a zooming concept map built with TheBrain technology (http://www.thebrain.com), and the movies will be indexed with subheads in a synchronised text window to allow ready access to relevant subtopics;
- students will be able to seek the advice of a tutor when they reach an impasse in their exercise discussions, an option not provided in the present implementation.

	Group							
Social and cognitive processes	D	E	F	G	J	K	M	SD
Nature of discussion and ownership of ideas	4	3	3	4	4	1	3.17	1.07
Amount of collaboration evident during session	4	3	3	4	4	1	3.17	1.07
Organisation of time during session	3	1	4	4	1	1	2.33	1.37
Building of understanding with BivarDescribe	1	2	3	3	2	1	2.00	0.82
Correction of statistical misunderstanding	0	1	4	2	2	0	1.50	1.38
Reference to demonstration movies	1	0	3	1	0	1	1.00	1.00
Mean	2.17	1.67	3.33	3.00	2.17	0.83		
Standard deviation	1.57	1.11	0.47	1.15	1.46	0.37		

Table 1: Ratings on six general social and cognitive process variables for groups participating in the video recording of the preparation session

	Process		Outcome	
	Mean	Range	Mean	Range
General understanding scales				
Level of understanding of statistical concepts	3.17	1 – 4	2.67	1 – 4
Integration of statistical concepts	2.33	0 – 4	2.33	1 – 4
Integration of levels of representation (graphical, conceptual, algebraic, computational)	2.50	0 – 4	2.33	1 – 4
Extent to which statistical concepts were related back to scenario	2.00	0 – 4	2.00	0 – 4
Sense of intended audience	2.17	0 – 4	2.17	0 – 4
Specific understanding scales				
Label and explain the variables correctly	2.33	0 – 4	2.83	1 – 4
Comment on the measurement scale of the variables	2.17	0 – 4	2.17	0 – 4
Note and explain the negative relationship between the Quest and Discrimination variables	1.83	0 – 4	1.83	0 – 4
Explain the strength of relationship using the PPV indicator and discuss why this relationship is strong or weak	2.67	1 – 4	2.83	2 – 4
Apply the variance change tool to the grand mean	1.83	1 – 4	2.67	0 – 4
Explain the changes to the data when applying the VCT the grand mean	1.67	0 – 3	2.33	0 – 4
Note and explain why LSC and PPV do not move when points move around grand mean with VC tool	2.17	0 – 4	2.17	0 – 4
Explain how variance change relative to grand mean influences interpretation of the scales	1.50	0 – 4	1.83	0 – 4

BEST COPY AVAILABLE

Table 2: Mean ratings (and obtained ranges) on six general and eight specific understanding scales for the preparation sessions (process) and teachback movies (outcome)

	Group							
Social and cognitive processes	D	E	F	G	J	K	M	**Range**
General understanding profiles	2.65	2.24	1.00	1.41	1.73	1.41	1.74	0 – 8.94
Specific understanding profiles	3.87	1.73	2.00	5.29	2.65	1.41	2.83	0 – 11.31

Table 3: Euclidean distances (dissimilarities) between the process and outcome profiles defined on the six general and eight specific understanding scales

References

Bain, J. D. (1999). Introduction: *Higher Education Research and Development, 18*, 165-172.

Becker, B. J. (1996). A look at the literature (and other resources) on teaching statistics. *Journal of Educational and Behavioral Statistics, 21*, 71-90.

Derry, S., Levin, J. R., & Schauble, L. (1995). Stimulating statistical thinking through situated simulations. *Teaching of Psychology, 22*, 51-57.

Fischer III, H. W. (1996). Teaching statistics from the user's perspective. *Teaching Sociology; 24* (2), 225-30.

Hays, W. L. (1981). *Statistics* (3rd Edition). NY: Holt, Rinehart and Winston.

Huberty, C. J. (1987). On statistical testing. *Educational Researcher, 16* (8), 4-9.

Jonassen, D. H., & Reeves, T. C. (1996). Learning with technology: Using computers as cognitive tools. In D. H. Jonassen, (Ed.), *Handbook of research on educational communications and technology* (pp. 693-719). NY: Macmillan.

Keppel, G., & Zedeck, S. (1989). *Data analysis for research designs: Analysis of variance and multiple regression/correlation approaches.* NY: Freeman.

Laurillard, D. (1993). *Rethinking university teaching: A framework for the effective use of educational technology.* London: Routledge.

McFarland, S. G., (1989). Religious orientation and targets of discrimination. *Journal for the Scientific Study of Religion, 28*, 324-336..

Pask, G. (1976). Conversational techniques in the study and practice of education. *British Journal of Educational Psychology, 46*, 12-25.

Perkins, D., & Blythe, T. (1994). Putting understanding up front. *Educational Leadership, 51*, 4-7.

Rosenthal, R., & Rosnow, R. L. (1985). *Contrast analysis: Focused comparisons in the Analysis of Variance.* NY: Cambridge University Press.

Roschelle, J. (1992). Learning by collaborating: Convergent conceptual change. *Journal of the Learning Sciences, 2*, 235-276.

Rosnow, R. L., & Rosenthal, R. (1989). Statistical procedures and the justification of knowledge in psychological science. *American Psychologist, 44*, 1276-1284.

Shaughnessy J. J., Zechmeister, E. B., & Zechmeister, J. S. (2000). *Research methods in psychology.* (5th Edn.) McGraw Hill: Boston.

Skemp, R. R. (1976). Relational understanding and instrumental understanding. *Mathematics Teaching, 77*, 20-26.

Thompson, B. (1993). *The general linear model (as opposed to the classical ordinary sums of squares) approach to analysis of variance should be taught to introductory statistical methods classes.* Paper presented at the annual meeting of the American Educational Research Association, Atlanta Georgia, April 12-16. ERIC Document ED 358 134.

Thompson, W. B. (1994). Making data analysis realistic: Incorporating research into statistics courses. *Teaching of Psychology, 21*, 41-43.

Wilkinson, L., & Task Force on Statistical Inference. (1999). Statistical methods in psychology journals. *American Psychologist, 54*, 594-604.

BEST COPY AVAILABLE

U.S. Department of Education
Office of Educational Research and Improvement (OERI)
National Library of Education (NLE)
Educational Resources Information Center (ERIC)

NOTICE

Reproduction Basis

[X] This document is covered by a signed "Reproduction Release (Blanket)" form (on file within the ERIC system), encompassing all or classes of documents from its source organization and, therefore, does not require a "Specific Document" Release form.

[ ] This document is Federally-funded, or carries its own permission to reproduce, or is otherwise in the public domain and, therefore, may be reproduced by ERIC without a signed Reproduction Release form (either "Specific Document" or "Blanket").